



REVIEW PAPER



Optimising sleep and performance during night float: A systematic review of evidence and implications for graduate medical education trainees

David Sholtes¹ | Howard M. Kravitz^{1,2} | Aniruddha Deka¹ | Jennifer Westrick³ | Louis F. Fogg⁴ | Michael Gottlieb⁵

¹Department of Psychiatry and Behavioral Sciences, Rush University Medical Center, Chicago, IL, USA

²Department of Preventive Medicine, Rush University Medical Center, Chicago, IL, USA

³Library of Rush University Medical Center, Rush University Medical Center, Chicago, IL, USA

⁴Department of Community, Systems and Mental Health Nursing, College of Nursing, Rush University Medical Center, Chicago, IL, USA

⁵Department of Emergency Medicine, Rush University Medical Center, Chicago, IL, USA

Correspondence

David Sholtes, Department of Psychiatry and Behavioral Sciences, Rush University Medical Center, 1653 W Congress Parkway, Chicago, IL 60612, USA.

Email: david_sholtes@rush.edu

Abstract

Graduate medical education (GME) training commonly requires residents and fellows to engage in night float shift work. This review aims to assess the effectiveness of interventions for trainees when preparing for, completing, and recovering from working night float shifts. We reviewed all available studies published prior to September 2019 using PubMed, Scopus, CINAHL, the Cochrane library, PsycINFO, and Google Scholar databases. We included all original, primary research articles assessing either non-pharmacological or pharmacological interventions on the chronobiological and physiological effects of night float shift work among GME trainees. Five studies ($n = 179$ patients) met inclusion criteria. Interventions included melatonin in the morning before sleep after night float shifts, napping during night float shifts, modafinil after a night of sleep deprivation, and caffeinated energy drinks after 6 consecutive night float shifts. Melatonin improved one measure of attention. A 2-hr nap was associated with improved speed related to task switching. Modafinil improved performance in tests of cognition. Caffeinated energy drinks led to improvement in select driving performance variables and reaction time. Effect sizes for outcome variables were calculated. Heterogeneity among the studies precluded combining the data in a meta-analysis. According to GRADE criteria, the quality of the evidence in these studies was low or very low. Our findings suggest GME trainees may benefit from utilising a limited number of interventions when preparing for or recovering from night float shift work. More investigation is needed to identify interventions that could help GME trainees adapt to and recover from working night float shifts.

KEYWORDS

ACGME, graduate medical education, night float, night shift, residency, shift work

1 | INTRODUCTION

Trainees (residents and fellows) in Graduate Medical Education (GME) training programmes often are required to complete night ('night float') shifts. The phrase 'night float' refers to a rotation or a

shift during which a GME trainee engages in clinical work during the evening and early morning and is off duty during the daytime before and after the shift. The duration and timing of night float shifts varies widely, and they differ from traditional extended duty overnight call shifts (i.e. working consecutively ≥ 24 hr). For example, night float

shifts may refer to a single night shift or multiple night shifts in a row, whereas extended duty overnight call shifts are scheduled up to every third night. The Accreditation Council for Graduate Medical Education (ACGME) defines 'night float' as:

A rotation or educational experience designed to either eliminate in-house call or to assist other residents during the night. Residents assigned to night float are assigned on-site duty during evening/night shifts and are responsible for admitting or cross-covering patients until morning and do not have daytime assignments. Rotation must have an educational focus.

(Accreditation Council for Graduate Medical Education Glossary of Terms, 2018)

The National Academy of Medicine defines "(day or night) float" as "a shift of residents that are not assigned to a single service but 'float' across services or teams to help with admissions and follow-up" (Ulmer et al., 2009).

To remain compliant with the 2003 ACGME duty hour requirements restricting GME trainees to working no more than 80 hr/week, GME training programmes in the United States have increasingly implemented night float rotations. For example, during the 2017–2018 academic year, the ACGME reported that 88% of accredited internal medicine programmes had a night float system (FREIDA Online, the AMA Residency, & Fellowship Database AMA, 2019), which was a significant increase compared with 30% in 1996 (Trontell et al., 1991) and 76% in 2006 (Wallach et al., 2006).

In general, night shift work is disharmonious with living habits and social activities. Night shift work is associated with increased risk of developing medical conditions such as metabolic syndrome, cardiovascular disease, and some cancers (Schernhammer et al., 2003; Wang et al., 2014; Wegrzyn et al., 2017). Furthermore, night shift work is associated with circadian rhythm misalignment; accumulated sleep debt; and may negatively impacting learning, performance, mood, and the ability to safely drive a motor vehicle (Costa, 2010). Maladaptation to night shift work can lead to "shift work disorder" (SWD), shift work type, also known as "Circadian rhythm sleep disorder" or CRSWD, shift work type, in the International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM). Summarised criteria for SWD from the third edition of the International Classification of Sleep Disorders (ICSD-3) are displayed in Table 1 (American Academy of Sleep Medicine, 2014).

To help GME trainees adapt to shift work, the ACGME requires training programmes to educate faculty members and residents on alertness management and fatigue mitigation processes, on how to recognise signs of fatigue and sleep deprivation, and to encourage the use of fatigue mitigation processes to manage the potential negative effects on patient care and learning (ACGME Common Program Requirements Section VI, 2017). Additionally, the 2017 ACGME Common Program Requirements Report outlines the following fatigue prevention strategies for GME trainees (ACGME Common Program Requirements Section VI, 2017):

TABLE 1 ICSD-3 diagnostic criteria for shift work disorder

ICSD-3 Diagnostic criteria for shift work disorder

- There is a report of insomnia and/or excessive sleepiness, accompanied by a reduction of total sleep time, which is associated with a recurring work schedule that overlaps with the usual time for sleep
- The symptoms have been present and associated with the shift work schedule for at least 3 months
- The symptoms cause clinically significant distress or impairment in mental, physical, social, occupational, education, or other important areas of functioning
- Sleep log and actigraphy monitoring (whenever possible and preferably with concurrent measurement of light exposure) for at least 14 days (work and free days) demonstrate a disturbed sleep and wake pattern
- The sleep and/or wake disturbance are not better explained by another current sleep disorder, medical or neurologic disorder, mental disorder, medication use, poor sleep hygiene, or substance use disorder

Strategies that may be used include, but are not limited to, strategic napping; the judicious use of caffeine; availability of other caregivers; time management to maximize sleep off-duty; learning to recognize the signs of fatigue, and self-monitoring performance and/or asking others to monitor performance; remaining active to promote alertness; maintaining a healthy diet; using relaxation techniques to fall asleep; maintaining a consistent sleep routine; exercising regularly; increasing sleep time before and after call; and ensuring sufficient sleep recovery periods.

However, the report does not include references to evidence supporting these recommendations nor does it provide GME trainees with practical ways to incorporate many of these strategies. Additionally, the report does not outline whether strategies may differ for night float shifts compared to 24–30 hr call shifts. Furthermore, specific to night float shifts, the ACGME duty hour changes in recent years have trended towards less restrictive duty hour requirements (e.g. maximum consecutive night float shifts no longer limited to 6), despite controversial evidence to support these changes, making it even harder for GME trainees to incorporate many of these strategies (ACGME Common Program Requirements (Residency), 2019; McHill et al., 2018).

The existing body of evidence available for the general population of night shift workers outlines a number of potentially helpful interventions for navigating night float work (Liira et al., 2014; McKenna & Wilkes, 2018; Neil-Sztramko et al., 2014; Slanger et al., 2016). However, these data may not generalise to GME trainees. As such, we cannot be certain that the conclusions drawn from these prior studies can be applied to the GME trainees who are a unique population of shift workers routinely encountering long work hours, sleep deprivation, and physical and emotional stress related to balancing their responsibility for patients.

Night float shift work differs from extended duty overnight call shifts and is commonplace in GME training (Ulmer et al., 2009; Weiss et al., 2016). In our literature search, we did not identify a review that focussed specifically on night work interventions among GME trainees. The objective of the present systematic review was to assess the effectiveness of interventions for mitigating the chronobiological and physiological impact of night float shifts that GME trainees experience.

2 | METHODS

2.1 | Literature search

A systematic search of the literature was conducted using PubMed, Scopus, Cumulative Index of Nursing and Allied Health Literature (CINAHL), the Cochrane CENTRAL Register of Controlled Clinical Trials, the Cochrane Database of Systematic Reviews, Google Scholar, and the PsycINFO databases. No limitations were set on language, date of publication, or geographical area. We found no relevant non-English papers. The search was conducted from the date of inception for each database through to 10 September 2019. The searches were performed by an experienced medical research librarian. The search focused on two main terms: “residents” and “night shift”. The search string for each of the databases is included in Appendix S1. In addition to the above search, we also reviewed the bibliographies of identified studies and review articles for potential missed articles, consulted with topic experts to help identify any further relevant studies, and searched the websites of the following well-established, national residency groups for relevant studies that might not be indexed in the biomedical databases listed above: Emergency Medicine Resident’s Association (EMRA), the Residents and Fellows section of the American Medical Association website (AMA), and Association of Family Medicine Residency Directors (AFMRD). We contacted authors of papers when there were uncertainties regarding study sample or design. The search conforms to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for systematic reviews and was performed in accordance with the best practice guideline (Liberati et al., 2009).

2.2 | Inclusion and exclusion criteria

Inclusion criteria included all original, primary research articles assessing either pharmacological or non-pharmacological interventions on the chronobiological and physiological effects of night float shifts among GME trainees (note that the term ‘GME trainees’ typically includes both residents and fellows, but papers included in our review identified only ‘residents’). We included all papers that were retrospective observational studies, prospective observational studies, or randomised controlled trials (RCTs), and outlined search terms are in Appendix S1. Exclusion criteria consisted of papers that:

(a) were not specific to residents or fellows; (b) did not investigate countermeasures (e.g. naps, melatonin, caffeine, bright light) or interventions (e.g. studied effects of night float shift only rather than comparing an intervention versus a control or placebo); (c) involved an intervention period that is not specific to night float shifts (e.g. 30-hr overnight call shifts); and (d) reviews, surveys, case reports, case series, and editorials. Using these criteria, two physician-investigators (DS, AD) independently assessed abstracts for eligibility. Abstracts that met initial criteria were reviewed as full manuscripts. Studies that met the eligibility criteria after full text review by both reviewers were included in the final data analysis. Disagreements were resolved by consensus, with the addition of a third reviewer if needed.

2.3 | Data extraction

The following characteristics were obtained for each of the included studies and were extracted by one investigator (DS) and verified for accuracy by a second investigator (HMK): last name of the first author, study title, publication year, total study population factors including size, gender, age, medical specialty, and years of training, and study design including intervention and outcomes.

2.4 | Bias assessment

The revised Cochrane Risk of Bias Tool for RCTs was used for the included studies (Higgins et al., 2011). A study was considered as having an overall low risk of bias (RoB) when there was low RoB in all of the following five domains: the randomisation process, deviation from intended interventions (i.e. effect of adhering to intervention or assignment to intervention), missing outcome data, measurement of the outcome, and selection of the reported results. The study was regarded as having an unclear RoB (e.g. ‘some concerns’) if at least one of these domains was unclear. The study was considered as having a high RoB if at least one of these domains showed a high RoB. Two reviewers (DS, HMK) independently assessed the evidence reported in the selected studies. Discrepancies in ratings were resolved by discussion with a third reviewer (MG).

2.5 | Evidence grading

The authors independently utilised the Grading of Recommendation, Assessment, Development, and Evaluation (GRADE) system via the GRADEpro tool (<https://gdt.gradeapro.org>) when evaluating the evidence related to individual outcomes as recommended by the Cochrane Collaboration (Ryan & Hill, 2016). The GRADE is a tool for rating quality of evidence for a specific intervention in research papers and for producing evidence-based recommendations for clinical practice (Ryan & Hill, 2016). The GRADE approach involves examining five methodological factors:

RoB, inconsistency, indirectness, imprecision, and publication bias. RCTs are initially categorised as high quality and can be downgraded based on factors associated with these five factors. The quality of evidence is then decided to be high, moderate, low, or very low. Two reviewers (DS, HMK) independently assessed the evidence reported in the selected studies and there were no disagreements on ratings.

2.6 | Data synthesis

Included studies were summarised using narrative synthesis. The studies were then quantitatively analysed with regard to the magnitude of effect sizes (ES) related to study outcome variables within studies (Becker, 2000). As ES were not determined in the original studies, we calculated Cohen's *d* (Cavallo et al., 2005; Huffmyer et al., 2019) or estimated Cohen's *d* outcomes using Microsoft Excel 2013 (Redmond, 2013) in collaboration with a statistician (LF). The estimated Cohen's *d* outcomes were calculated using two-tailed Student's *t* distribution (Jockovich et al., 2000) or *F* statistic (Huffmyer et al., 2019; Sugden et al., 2012; Tempesta et al., 2013). A more detailed description of the calculations is included in the Appendix S2. Data were not combined for meta-analysis due to heterogeneous study design and outcome measures among the included reports.

3 | RESULTS

3.1 | Characteristics of included studies

A total of 2,818 citations were identified via the following electronic databases with total number of citations for each database listed in parentheses: PubMed (1,030), Scopus (1,057), CINAHL (334), Cochrane CENTRAL Register of Controlled Clinical Trials (86), Cochrane Database of Systematic Reviews (1), PsychINFO (207), and Google Scholar (100). The Google Scholar search was limited to the first 100 citations as recommended by Bramer et al. (2017). After duplicates were removed, 1,635 articles remained. In all, 42 articles were reviewed as full-text manuscripts and five papers ($n = 179$ participants) were selected for final inclusion (Figure 1). The five selected studies include three RCTs (Huffmyer et al., 2019; Jockovich et al., 2000; Tempesta et al., 2013), one prospective concurrent quasi-experimental trial (Sugden et al., 2012), and a prospective randomised double-blind crossover study (Cavallo et al., 2005). Of the 179 participants, 103 were male (58%). All studies included residents working at university-affiliated hospitals in specialties including emergency medicine (EM) (Jockovich et al., 2000), paediatrics (Cavallo et al., 2005), anaesthesiology (Huffmyer et al., 2019), and surgery (Tempesta et al., 2013). Three studies were conducted in the United States (Cavallo et al., 2005; Huffmyer et al., 2019; Jockovich et al., 2000), one study was conducted in Italy (Tempesta et al., 2013),

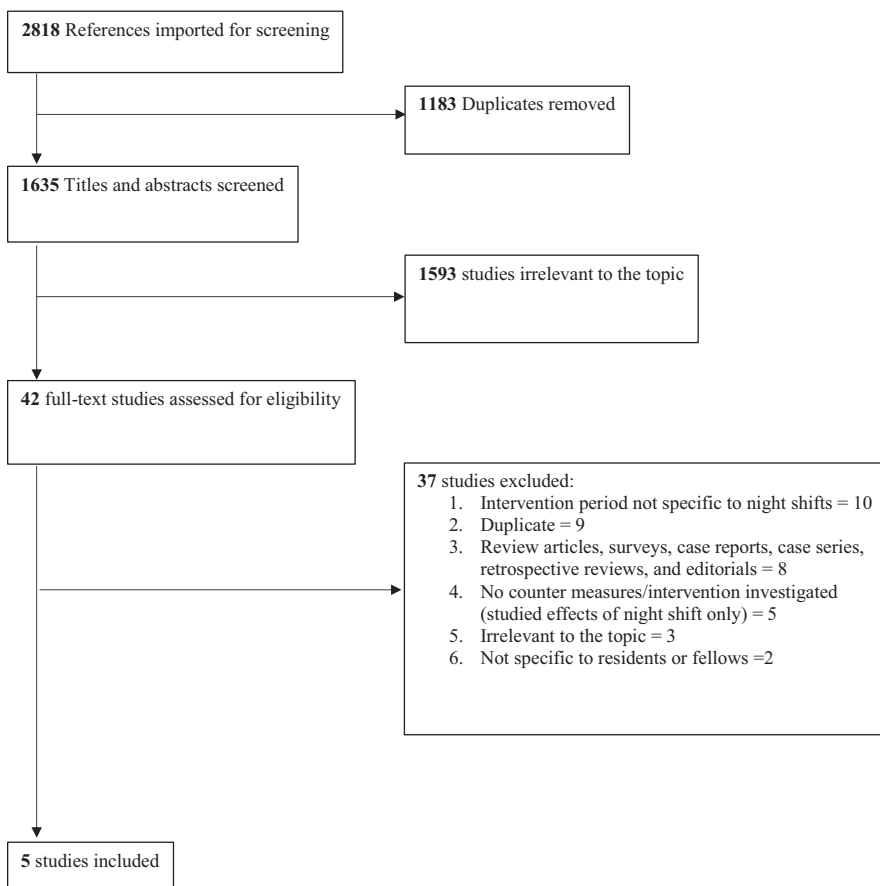


FIGURE 1 PRISMA flow diagram

and one study was conducted in the UK (Sugden et al., 2012). Study design and sample characteristics are shown in Table 2.

3.2 | Main findings

The following interventions were investigated: (a) melatonin in the morning before sleep after working night float shifts (Cavallo et al., 2005; Jockovich et al., 2000), (b) napping during night float shifts (Tempesta et al., 2013), (c) modafinil after a night of sleep deprivation (Sugden et al., 2012), and (d) using a caffeinated energy drink after night float shift (Huffmyer et al., 2019) (Table 2). ES were calculated for all outcome variables (Table 2).

Jockovich et al. (2000) gave 19 EM residents a 1 mg oral dose of melatonin in the morning after 3 consecutive night float shifts and measured self-reported mood (Profile of Mood State, POMS) and alertness (Stanford Sleepiness Scale [SSS]) in the evening, prior to night float shifts, and total sleep duration during recovery sleep (actigraphy). They found no beneficial effects of the 1-mg melatonin dose, compared to placebo, on recovery sleep ($ES = 0.2$), alertness ($ES = 0.12$), or mood state ($ES = 0.12$) during night float shift work among EM residents.

Cavallo et al. (2005) tested whether a 3 mg oral dose of melatonin given to 45 paediatric medicine residents in the morning after night float shifts improved sleep duration (sleep diary) during recovery sleep, mood (POMS), and attentional related problems (Conners' Continuous Performance Test [CPT]). The residents were assessed for mood and attention in the morning, after night float shifts. Results showed that this larger melatonin dose, compared to placebo, significantly improved one measure of attention, number of omission errors ($ES = 0.11$), but did not demonstrate a statistically significant difference in recovery sleep duration, mood, or other measures of attention.

The nap study (Tempesta et al., 2013) compared the effect of a 2-hr nap (nap group, NG, $n = 16$) during a night float shift with two non-napping groups. Performance of the NG on executive functioning skills (e.g. task-switching and go/no go tasks) was compared with performance following either no naps in a wake group (WG, $n = 16$) or in a normal sleep at home (sleep group, SG, $n = 22$). The investigators found that the NG and SG participants demonstrated improvement in performance speed related to task switching relative to baseline. Using the available outcome data, the ES of 3.38 was calculated by converting an F score of 2.86 associated with mixed-model analysis of variance (ANOVA) using 'group' (SG, WG and NG) as the between-subject factor and 'session' (days 1, 2 or 3) as the within-subject factor.

Sugden et al. (2012) examined whether taking a 200 mg oral dose of modafinil given to 20 residents after a night of sleep deprivation enhanced cognitive performance (Cambridge Neuropsychological Test Automated Battery [CANTAB]), psychomotor performance (Minimally Invasive Surgical Trainer Virtual Reality [MIST-VR]), and subjective feelings such as alert-drowsy, attentive-dreamy, and incompetent-proficient (visual analogue scales). Compared to placebo

($n = 19$), modafinil improved cognitive performance specific to working memory ($ES = 4.6$), flexibility to redirect attention ($ES = 4.3$), spatial planning ($ES = 4.2$), and impulsive decision making ($ES = 5.2$), but not psychomotor performance under time pressure.

Finally, the Huffmyer study (Huffmyer et al., 2019) used a cross-over design to investigate whether ingesting a caffeinated energy drink with 160 mg of caffeine in the morning compared with a non-caffeinated drink (single blind). A total of 22 anaesthesiology residents completed 6 consecutive night-float shifts improved driving performance (simulated driving using the Driver Guidance System) and reaction time (Psychomotor Vigilance Test). Compared to the non-caffeinated energy drink, the caffeinated drink was associated with temporally mixed results on the driving simulation that were observed 60 min after ingestion. Residents' demonstrated poorer driving performance during the first 10 min of driving (Epoch 1, open road segment) related to control of steering, speed, throttle, and number of collisions. During the subsequent 30 min of driving (Epochs 2 and 3, open road and obstacle segments), the caffeinated energy drink was associated with improved performance on lane position in both open road ($ES = 2.9$) and obstacle segments ($ES = 1.3$) during the last 15 min of driving (Epoch 3). Additionally, it was associated with fewer collisions ($ES = 1.3$) during the last 30 min of driving (Epochs 2 and 3), and less deviation in speed ($ES = 1.7$) in the last 15 min of driving during the obstacle segment (Epoch 3). The caffeinated energy drink group had an improved mean reaction time by 15.1 ms ($ES = 0.46$), but did not have meaningful impact on subjective sleepiness reports.

3.3 | Risk of bias (RoB) assessment

The RoB of the included studies ranged from low to high (Figure 2). The studies had low (Sugden et al., 2012), some, (Jockovich et al., 2000) or high RoB (Cavallo et al., 2005; Huffmyer et al., 2019; Tempesta et al., 2013) due to insufficient information reported about allocation concealment (Cavallo et al., 2005; Huffmyer et al., 2019; Jockovich et al., 2000; Tempesta et al., 2013), blinding (Cavallo et al., 2005; Huffmyer et al., 2019; Jockovich et al., 2000; Sugden et al., 2012), randomisation (Cavallo et al., 2005; Jockovich et al., 2000; Tempesta et al., 2013), deviations from intended interventions (Huffmyer et al., 2019; Jockovich et al., 2000; Tempesta et al., 2013), missing outcome data (Cavallo et al., 2005; Tempesta et al., 2013), measurement of the outcome (Huffmyer et al., 2019; Jockovich et al., 2000; Tempesta et al., 2013), and selection of the reported results (Huffmyer et al., 2019; Jockovich et al., 2000; Tempesta et al., 2013). The nap study by Tempesta et al. (2013) had the overall highest RoB, due to group assignment not being randomised and non-equivalence of groups because the design was modified from a randomised trial to a quasi-experimental study, with composition of "intervention" groups depending on ward-based demands (nap group) and residents' individual choices (wake group), missing outcome data, and lack of control for confounding and selection bias.

TABLE 2 Main findings

Author (year)	Study design	Study sample	Sample size	Average age, years	Gender	Intervention	Outcome measures	Outcomes	Effect size (ES)
Tempesta et al. (2013)	Prospective, concurrent quasi-experimental trial (ecological validity)	Surgery interns conducted at a university affiliated hospital in Italy	N = 54 NG, n = 21 WG, n = 11	29	17 Males, 37 Females	2-hr nap versus no nap versus sleep at home (control) group during night shift	1. Effect of long (ad libitum) nocturnal nap (i.e. ~2 hr) during night shift on executive function skills via task switching and go/no go task 2. Actiographic TST 3. Sleepiness via KSS 4. POMS	Sleep group and NG participants with better task switching speed compared to WG participants. Napping associated with improvement in performance speed relative to baseline, but <effect of regular night's sleep, and "skill specific" (task switching, not go/no go task). Comparable levels of performance were achieved by both the sleep and nap groups For actiographic TST, WG slept longer during recovery Similar results for subjective outcome measures	Comparing all three groups over time Task switching speed = 3.4 TST = 18.5 KSS = 5.4 POMS Total = 4.9 Comparing all three groups to type of trial over time Task switching accuracy = 5.2 Go/no go accuracy = 3.2
Cavallo et al. (2005)	Double-blind, randomised, placebo-controlled crossover	Second-year paediatric residents working night shifts at a university affiliated tertiary-care paediatric hospital in the USA	N = 45 Melatonin and placebo, n = 28 Melatonin only, n = 7 Placebo only, n = 10	28.6	16 males, 29 Females	Melatonin 3 mg versus placebo in morning before sleep after working night shifts	1. Subjective sleep measures including sleep diary for total sleep, sleep latency and VAS for sleep quality 2. Subjective fatigue via POMS 3. Assessment of attention-related problems via Conners' CPT	One measure of attention, the number of omission errors, was significantly lower on melatonin than on placebo No statistically significant difference in measures of sleep, mood, and 5 of 6 measures of attention otherwise	Total SID d = 0.10 POMS total = 0.3 Conners' CPT Number of commission errors = 0.03 Number of omission errors = 0.11 Mean reaction time = 0.10 Hit reaction time block change = 0.03 Attentiveness = 0.1 Risk taking = 0

(Continues)

TABLE 2 (Continued)

Author (year)	Study design	Study sample	Sample size	Average age, years	Gender	Intervention	Outcome measures	Outcomes	Effect size (ES)
Sugden et al. (2012)	Parallel, double-blind, randomised, and placebo-controlled study	Third year residents working at an academic medical centre in London, UK. Residency training speciality information not available	N = 39 Modafinil, n = 20 Placebo, n = 19	28	39 males, 0 females	Modafinil 200 mg versus lactose placebo after 1 night of sleep deprivation	1. CANTAB neuropsychological tests to assess higher cognitive function 2. MIST-VR to assess clinical psychomotor performance 3. Subjective feelings via VAS 4. Safety measures	Modafinil improved cognitive performance processes critical for efficient information processing, flexible thinking, and decision-making under time pressure, but not psychomotor performance	Cognitive Tasks IED: Stages completed = 4.3 OTS: Mean latency = 3.1 Mean latency to correct, 5 move problems = 4.2 Mean attempts = 0.4 CGT: Probability of choosing the most likely outcome = 1.5 Percentage bet placed on decision = 1.2 Overall proportional bet = 0.2 Deliberation time = 1.2 Delay aversion = 5.2 Reverse SSP: Errors = 4.6 Span length = 3.2 Psychomotor Tasks Manipulate diathermy: Mean time taken = 1.4 Mean error rate = 0.4 Mean economy of movement = 1.1 No. of attempts satisfying all benchmarked performance criteria = 0.03 Stretch diathermy: Mean time taken = 2.2 Mean error rate = 0.3 Mean economy of movement = 1.8 No. of attempts satisfying all benchmarked performance criteria = 0.3 Subjective feelings Alertness = 5.1 Strength = 6.0 Proficiency = 3.7 Gregariousness = 4.7 Attentiveness = 6.1 Safety measures HR = 0.3 SBP = 1.6 DBP = 0.9

(Continues)

TABLE 2 (Continued)

Author (year)	Study design	Study sample	Sample size	Average age, years	Gender	Intervention	Outcome measures	Outcomes	Effect size (ES)
Jockovich et al. (2000)	Prospective, randomised, double-blind crossover study	19 EM residents from the USA in urban ED who worked two series of at least 3 consecutive nights separated by at least 1 week of day shifts	N = 19	28.2	15 males, 4 females	Melatonin 1 mg versus placebo morning after working night shifts	1. TSD, SIL, SIE using the Actigraph 1,000 2. Subjective fatigue via the POMS 3. SSS	Melatonin did not improve sleep quality (SIE, SID, SIL), alertness/sleepiness, or mood	TSD = 0.2 POMS total = 0.02 SSS = 0.12
Huffmyer et al. (2019)	Prospective, randomised, single-blind crossover study (residents blinded)	26 anaesthesiology residents (PGY 2–4) from an academic medical centre in the USA who worked 6 consecutive night shifts separated by at least 2–4 months prior to crossover	N = 22 included in analysis	30	16 males, 6 females	Caffeinated energy drink (160 mg caffeine) versus non-caffeinated energy drink after 6 consecutive night shifts	1. Primary outcomes: High-fidelity driving simulation performance—lane position, collisions, off-course driving 1 hr after consuming drink in both open road and collision segments 2. Additional outcomes: Reaction time and lapses in attention on PVT 3. Additional measures include ESS and SDQ	Caffeinated energy drink led to initially poorer control of driving variables during Epoch 1, improved performance on lane position in open road during Epoch 3, significantly fewer collisions in Epochs 2 and 3, less deviation in speed during obstacle segment in epoch 3. Caffeinated energy drink significantly improved mean reaction time (15.1 ms). No statistically significant differences in self-reported sleep over the final night-float shift for the week, average sleep for the 6-shift night-float period, caffeine sensitivity, ESS or SDQ	Open road segment Throttle = 3.9 Speed = 3.8 Steering = 4.3 Position = 2.9 Obstacle segment Throttle = 2.2 Speed = 1.7 Steering = 1.4 Position = 2.3 Obstacles = 1.3 PVT Mean reaction time = 0.46 Major lapses = 0.36 Minor lapses = 0.42 ESS pre-drive: 0.01 ESS post-drive: 0.08 SDQ pre-drive/post-drive How safe to drive a car now = 0.02/−0.02 How alert are you now = 0.01/−0.03 How sleepy are you now = 0.01/0.02

Abbreviations: CANTAB, Cambridge Neuropsychological Test Automated Battery; CGT, Cambridge gamble task; Conners' CPT, Conners' Continuous Performance Test; DBP, diastolic blood pressure; EM, emergency medicine; ESS, Epworth Sleepiness Scale; HR, heart rate; IED, intra-extra-dimensional set shift; KSS, Karolinska Sleepiness Scale; MIST-VR, Minimally Invasive Surgical Trainer Virtual Reality; NG, nap group; OTS, One-Touch Stockings of Cambridge; POMS, Profile of Mood States; PVT, Psychomotor Vigilance Task; SBP, systolic blood pressure; SID, sleep duration; SDQ, Sleepiness and Driving Questionnaire; SIE, sleep efficiency; SIL, sleep latency; SSP, spatial span; SSS, Stanford Sleepiness Scale; TSD, total sleep duration; TST, total sleep time; VAS, visual analogue scale; WG, wake group (no nap).

FIGURE 2 Risk of bias according to revised Cochrane Risk of Bias Tool for RCTs (RoB 2.0)

Reference	Randomization process	Deviations from intended interventions	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall	
Huffmyer et al, 2019	?	?	+	-	+	-	+
Tempesta et al, 2013	-	-	-	-	?	-	?
Cavallo et al, 2005	?	?	-	+	+	-	-
Jockovich et al, 2000	?	?	+	?	?	?	
Sugden et al, 2012	+	+	+	+	+	+	

Low risk
Some concerns
High risk

One co-author of the modafinil paper reported a potential conflict of interest (Sugden et al., 2012). Some authors did not state whether they had any potential conflict of interest (Cavallo et al., 2005; Jockovich et al., 2000), while others did explicitly report having no conflicts of interest (Huffmyer et al., 2019; Sugden et al., 2012; Tempesta et al., 2013). The Jockovich group (Jockovich et al., 2000) did not report sources of funding for the research, while the funding for all other studies (Cavallo et al., 2005; Huffmyer et al., 2019; Tempesta et al., 2013) came from institutional or federal grants.

3.4 | Quality of the evidence

A summary of the evidence quality according to the GRADE system is shown in Table 3. Indirectness of the evidence impacted the quality of the studies due to differences in study populations (e.g. surgery trainees have different demands and scheduling requirements than paediatric or EM trainees). Another source of variability is that two studies were conducted at sites outside of the United States (Sugden et al., 2012; Tempesta et al., 2013), where duty hour policies may differ from ACGME policies followed by trainees at the three United States study sites (Cavallo et al., 2003; Huffmyer et al., 2019; Jockovich et al., 2000). Serious inconsistencies in the results were attributed to differences in interventions and study design issues, such as variability in doses of melatonin (Cavallo et al., 2005; Jockovich et al., 2000), lengths of night float shifts (Cavallo et al., 2005; Huffmyer et al., 2019; Jockovich et al., 2000; Sugden et al., 2012; Tempesta et al., 2013), how sleep changes were measured (Cavallo et al., 2005; Jockovich et al., 2000; Tempesta et al., 2013), timing of interventions (Cavallo et al., 2005; Huffmyer et al., 2019; Jockovich et al., 2000; Sugden et al., 2012; Tempesta et al., 2013) when measures were completed (Cavallo et al., 2005; Huffmyer et al., 2019; Jockovich et al., 2000; Sugden et al., 2012; Tempesta et al., 2013), and what results were reported or available (Cavallo et al., 2005; Tempesta et al., 2013). All studies suffered from imprecision due to small sample sizes and, in the melatonin studies (Cavallo et al., 2005; Jockovich et al., 2000), this issue magnifies the detrimental effects of potential noncompliance with medication

dosing. Additionally, publication bias could not be assessed due to the limited number of studies.

4 | DISCUSSION

To our knowledge, this is the first systematic review and synthesis of the available evidence related to assessments of either non-pharmacological or pharmacological interventions on the chronobiological and physiological effects of night float shift work among GME trainees. The present review includes data from five studies and a total of 179 participants. The interventions included melatonin in the morning before sleep after working night float shifts, napping during night float shifts, modafinil after a night of sleep deprivation, and caffeinated energy drink after 6 consecutive night float shifts. These studies demonstrated that: (a) melatonin improved one measure of attention; (b) a 2-hr nap during a night float shift was associated with improvement in performance speed related to task switching relative to baseline compared to no nap; (c) modafinil after a night of sleep deprivation improved cognitive performance in some domains, but did not improve psychomotor performance; and (d) caffeinated energy drink consumption the morning after completing 6 consecutive night float shifts led to initial worsening of simulated driving performance (despite improvement in psychomotor vigilance) followed by some improvement of performance and improved reaction time.

Two studies excluded from the present systematic review had heterogeneous or 'mixed' study samples (i.e. participants were not limited to GME trainees), but merit discussion in light of our findings. One study was a double-blind randomised, placebo-controlled crossover trial that included EM attending physicians and residents who were assigned to take either melatonin 10 mg or placebo the morning after working night float shifts to study its effects on sleep quality (Farahmand et al., 2018). In this study, melatonin had limited benefit on sleep quality, similar to the melatonin studies that were included in the systematic review. Moreover, this excluded

TABLE 3 GRADE assessment of quality of included studies

Quality assessment									
Outcomes	Subjects, n	Included RCTs, n	RoB	Inconsistency	Indirectness	Imprecision	Publication bias	Other	Overall quality of evidence
Caffeinated energy drink versus non-caffeinated energy drink after 6 consecutive night float shifts									
Driving simulation performance	26	1 RCT (Huffmyer et al., 2019)	Serious RoB ^{a,b}	Serious ^c	No serious indirectness	Serious ^d	Unable to assess ^e	Intervention assesses use of energy drinks with many active ingredients confounds effect of caffeine (e.g. caffeinated drink has guarana extract, also a stimulant; both versions have taurine, panax ginseng extract, L-carnitine, L-arginine, glucuronolactone, inositol, maltodextrin), four participants used sleep aids during night float week which could also confound outcome data	Low
Reaction time and lapses in attention (assessed with: PVT)	26	1 RCT (Huffmyer et al., 2019)	Serious RoB ^{a,b}	Serious ^c	No serious indirectness	Serious ^d	Unable to assess ^e		Low
Melatonin compared to placebo for recovery sleep after night float shifts									
Sleep measures (assessed with: Sleep diary, actigraphy)	64	2 RCT (Cavallo et al., 2005; Jockovich et al., 2000)	Unclear RoB ^{b,i,g}	Serious ^c	No serious indirectness	Serious ^d	Unable to assess ^e	Study designs did not control for many factors that can influence outcomes including pharmacogenetic characteristics, individual tolerance to shift work, and a relationship of response to melatonin with individual tolerance to shift work	Very low
POMS	64	2 RCT (Jockovich et al., 2000 and Cavallo et al., 2005)	Unclear RoB ^{b,i,g,h}	Serious ^c	No serious indirectness	Serious ^d	Unable to assess ^e		Very low
Subjective sleepiness (assessed with: SSS)	19	1 RCT (Jockovich et al., 2000)	Unclear RoB ^{b,i,g}	Serious ^c	No serious indirectness	Serious ^d	Unable to assess ^e		Very low
Assessment of attention related problems (assessed with: Conners' CPT)	45	1 RCT (Cavallo et al., 2005)	Unclear RoB ^{b,i,g}	Serious ^c	No serious indirectness	Serious ^d	Unable to assess ^e		Very low

(Continues)

TABLE 3 (Continued)

Quality assessment										Overall quality of evidence
Outcomes	Subjects, n	Included RCTs, n	RoB	Inconsistency	Indirectness	Imprecision	Publication bias	Other		
2-hr nap compared to no nap versus sleep at home (control) for during night float shifts										
Executive function skills via task switching and go/no go task	54	1 RCT (Tempesta et al., 2013)	Serious RoB ^{b,i}	Serious ^c	No serious indirectness	Serious ^d	Unable to assess ^e	3 groups non-equivalent, composition of "intervention" groups depended on ward-based demands (NG) and interns' individual choices (WG) – i.e., investigators did not have full control of intervention		Very low
Actiographic TST	54	1 RCT (Tempesta et al., 2013)	Serious RoB ^{b,i}	Serious ^c	No serious indirectness	Serious ^d	Unable to assess ^e			Very low
Sleepiness (assessed with: KSS)	54	1 RCT (Tempesta et al., 2013)	Serious RoB ^{b,i}	Serious ^c	No serious indirectness	Serious ^d	Unable to assess ^e			Very low
Mood (assessed with: POMS)	54	1 RCT (Tempesta et al., 2013)	Serious RoB ^{b,i}	Serious ^c	No serious indirectness	Serious ^d	Unable to assess ^e			Very low
Modafinil compared to placebo for cognitive or psychomotor performance after 1 night of sleep deprivation										
Cognitive function (assessed with: CANTAB)	39	1 RCT (Sugden et al., 2012)	No serious RoB ^b	Serious ^c	No serious indirectness	Serious ^d	Unable to assess ^e	Data do not address effect of enhancement in older participants, or participants with poor baseline performance and using tasks that require the integration of cognitive and psychomotor skill. Authors' caveat: "Until more detailed laboratory and field evidence is accrued, the use of pharmacological agents to enhance performance in the workplace cannot be recommended".		Low
Sleepiness (assessed with: KSS)	39	1 RCT (Sugden et al., 2012)	No serious RoB ^b	Serious ^c	No serious indirectness	Serious ^d	Unable to assess ^e			Low

Abbreviations: CANTAB, Cambridge Neuropsychological Test Automated Battery; Conners' CPT, Conners' Continuous Performance Test; KSS, Karolinska Sleepiness Scale; POMS, Profile of Mood States; PVT, Psychomotor Vigilance Test; RCT, randomised controlled trial; RoB, risk of bias; SSS, Stanford Sleepiness Scale; WG, wake group (no nap).

^aMostly due to measurement of outcome, cross-over design, used repeated measures analysis but did not report whether there was a significant order effect or interactions.

^bInsufficient information reported about whether participants/study staff aware of assigned intervention during the trial.

^cInconsistency mainly due to differences in intervention, details of intervention, and methodological quality among these studies.

^dSmall sample size.

^eIt was not possible to check publication bias because of the limited number of trials for this outcome.

^fDue to cross-over design.

^gInsufficient information reported about randomisation process to determine whether random allocation sequence.

^hSome outcome data not available.

ⁱDue to non-random assignment based on demands and intern choice leading to potential self-selection bias and deviations from intended interventions.

study provided very low quality evidence that using melatonin has a beneficial impact on sleep duration, mood, and attention related problems.

Another randomised, controlled intervention trial investigated whether a scheduled 40-min nap improved the cognitive and psychomotor performance in EM residents and registered nurses working 12-hr night shifts (Smith-Coggins et al., 2006). The investigators reported that the nap intervention had beneficial effects on physician and nurse performance specific to reaction times, subjective feeling of vigour and fatigue, and memory, but did not improve performance on a simulated drive home after the night shift.

The data from these studies specific to GME trainees would add to the overall findings from our review. We were unable to retrieve summary data from studies by contacting the corresponding authors, so we could not analyse the data specific to residents. As a result, we did not include these two studies in the systematic review, but have discussed these findings.

4.1 | Implications for GME training

These have significant implications for GME trainees, GME administrators including programme directors, and accreditation bodies (e.g. ACGME). The large ES from the modafinil (Sugden et al., 2012) and caffeinated energy drink study (Huffmyer et al., 2019) may have the most significant implications for GME trainees and administrators. For example, in the context of engaging in cognitively demanding tasks after a night of sleep deprivation, modafinil may be of benefit. Additionally, after completing 6 night float shifts, the consumption of a caffeinated energy drink may impact driving performance and reaction time as outlined previously. However, the use of pharmacological enhancement in the context of sleep deprivation may contribute to additional circadian rhythm misalignment (Czeisler, 2010; Czeisler et al., 2009; Rose & Curry, 2009). The ES for the melatonin studies were relatively small (i.e. 0–0.3) implying limited benefit for GME trainees to use melatonin after night float shifts (Shy et al., 2011). Importantly, and a strength of the present paper, the findings identify knowledge gaps that may serve to improve transparency, emphasise empirical evidence, highlight inconsistencies and flawed study designs, and suggest areas for future research. Other strengths include using a protocol for the search and review process as outlined in the methods section. The paper also followed the GRADE approach towards assessing quality and the Revised Cochrane Risk of Bias Tool for RCTs, as recommended by the Cochrane collaboration (Higgins et al., 2011; Ryan & Hill, 2016; Sterne et al., 2019).

Despite the pervasiveness of night float shifts and the ACGME requirement that programmes educate GME trainees about alertness management and fatigue mitigation, medical trainees are lacking sleep medicine knowledge and educational resources (Kirsch & Khosla, 2019). A 2011 survey of United States and Canadian medical schools indicated that medical students receive an average of 187 min of sleep education during medical school training (Mindell et al., 2011). Another 2013 survey showed that the majority of UK-based residents who completed

night shifts were not aware of basic concepts related to increasing alertness and fatigue prevention (e.g. taking a prophylactic nap prior to a night shift, the impact of sleep inertia on alertness, utilising interventions for adapting to the circadian rhythm) (Jackson & Moreton, 2013). No such survey data could be found for United States GME trainees.

4.2 | Limitations

It is important to consider several additional limitations to the included studies. The heterogeneity in the interventions, environments, and subjects as well as non-comparability of ES precluded the ability to perform meta-analysis of the data, which is a major limitation. The ES calculations for the included studies ranged widely (0–18.5) and only one study reported the primary outcome measure (Huffmyer et al., 2019). Due to limitations in the data that were available, the ES calculations for the Tempesta paper either compare all three groups (nap, no nap, sleep at home) over time (days 1, 2, 3) or type of trial (task switching, go/no-go task, Karolinska Sleepiness Scale [KSS], actigraphy) over time leading to relatively large ES (range 3.4–18.5). For example, for total sleep time the ES calculation of 18.5 is not practically useful, as it is comparing WG, NG, and SG. It may be problematic to draw conclusions from the Tempesta paper about whether napping during night float shifts leads to improvements in executive functioning, alertness, or sleep because outcome data are missing. Furthermore, the Tempesta paper initially appeared to be a clinical trial, but was redesigned as a quasi-experimental study (i.e. non-random assignment, no control over intervention, and non-equivalent comparison groups) based on training demands and intern choice, potentially allowing serious deviations from the intended intervention and bias due to participants' expectations of benefits from their choice preferences. The Cavallo, Jockovich, and Huffmyer protocols were limited by their cross-over designs, which did not include analysis for potential time or intervention order effects. The Huffmyer study was also limited by having a single rather than a double-blind design (i.e. residents blinded to contents of drink). None of these studies examined the effects of circadian resetting with bright light therapy before and after shift changes, an intervention with demonstrated benefits in SWD.

Additionally, it is unclear how the findings from each specialty training programme (e.g. paediatrics, EM, anaesthesiology, and surgery) would generalise to other GME programmes, including fellowship training, which may have different programmatic demands. Moreover, due to the strict protocol design for most studies, the investigators' findings may have limited practical applicability for programme directors and would have unclear implications for GME trainees. For example, trainees considering ingesting a caffeinated energy drink prior to driving home after completing 3 or 7 night float shifts, as opposed to the 6 used in the study, would appreciate more clear evidence-based guidance regarding options. Other significant factors were not controlled for in these studies, such as the influence of chronic sleep deprivation, pharmacogenetic characteristics, individual tolerance to shift work, length of night float shifts, influences of activity, caffeine

use, and ambient/environmental factors such as exposures to bright light, temperature, and noise. Finally, although an experienced medical librarian (JW) performed a comprehensive search, and we used best practice methodology, as well as included several relevant grey literature (Paez, 2017) sources (e.g., theses, dissertations, conference posters and papers, ongoing research, and committee and government reports, as outlined in the literature search sub-section), it is possible that we may have missed potentially relevant papers.

5 | CONCLUSION

Graduate medical education training programmes are complex systems (Plsek, 2001), and modifications to ACGME duty hours regulations should be based on empirical evidence that considers the impact on trainees well-being and education, as well as on patient quality of care and safety (Fletcher et al., 2010; Rosenbaum & Lamas, 2012). With these goals in mind, the present review has assessed and summarised the available published data on interventions that may be beneficial for GME trainees to utilise when preparing for or recovering from night float shift work. The study outcomes suggest GME trainees may benefit from using interventions during/after night float shifts. However, the quality of this evidence is considered to be low or very low, highlighting the need for further investigation. Our present analysis was limited due to a paucity of available data. Without good data, it is not surprising that there is insufficient evidence for interventions to guide graduate training programmes to help GME trainees navigate working night float shifts. Future studies involving trainees from a broader and more diverse range of medical and surgical specialties are needed.

DISCLOSURE STATEMENT

Work for this study was completed at Rush University Medical Center in Chicago, IL. All authors have seen and approved the manuscript. Dr Sholtes is a primary investigator for a study funded by LivaNova USA, investigating safety and effectiveness of VNS therapy system as adjunctive therapy in subjects with treatment resistant depression, a sub-investigator for a study funded by Roche investigating the efficacy, and safety of balovaptan in adults with autism spectrum disorder, and for a study funded by Janssen investigating the efficacy, and safety of esketamine in adults with major depressive disorder. Dr Kravitz receives research grant funding from the National Institutes of Health (NIH)/ National Institute on Aging, and has had support from The Stanley G. Harris Family Chair of Psychiatry. The other authors report no conflicts of interest. No authors received funding or financial support for the development of this manuscript.

AUTHOR CONTRIBUTIONS

All authors have contributed to this article. DS, HMK, and MG conceived and designed this research. JW, DS, and AD searched the databases, extracted the data, and screened the trials. DS and HMK appraised the quality of the included SRs and RCTs. DS, LF, HMK, and MG performed the analysis and interpretation of data. DS, MG, and HMK drafted the

full text. DS, MG, and HMK critically revised the manuscript for important intellectual content. All authors reviewed the manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

David Sholtes  <https://orcid.org/0000-0003-2230-7436>

Michael Gottlieb  <https://orcid.org/0000-0003-3276-8375>

REFERENCES

- Accreditation Council for Graduate Medical Education (2017). *ACGME Common Program Requirements Section VI with Background and Intent*. Retrieved from https://www.acgme.org/Portals/0/PFAAssets/ProgramRequirements/CPRs_Section%20VI_with-Background-and-Intent_2017-01.pdf
- Accreditation Council for Graduate Medical Education (2018). *Accreditation council for graduate medical education glossary of terms*. Retrieved from <http://tagme.org/wp-content/uploads/2019/04/ACGME-Glossary-of-Terms-2018.pdf>. Published May 8, 2019.
- Accreditation Council for Graduate Medical Education (2019). *ACGME Common Program Requirements (Residency)*. Retrieved from <https://www.acgme.org/Portals/0/PFAAssets/ProgramRequirements/CPRResidency2019.pdf>
- American Academy of Sleep Medicine (2014). *International classification of sleep disorders (3rd ed.)*. Darien, IL: American Academy of Sleep Medicine.
- Becker, L. A. (2000). *Effect size (ES)*. Retrieved from <http://web.uccs.edu/lbecker/Psy590/es.htm>. Published March 21, 2000.
- Bramer, W. M., Rethlefsen, M. L., Kleijnen, J., & Franco, O. H. (2017). Optimal database combinations for literature searches in systematic reviews: A prospective exploratory study. *Systematic Reviews*, 6, 245. <https://doi.org/10.1186/s13643-017-0644-y>
- Cavallo, A., Ris, M. D., & Succop, P. (2003). The night float paradigm to decrease sleep deprivation: Good solution or a new problem? *Ergonomics*, 46(7), 653–663. <https://doi.org/10.1080/001401303100085671>
- Cavallo, A., Ris, M. D., Succop, P., & Jaskiewicz, J. (2005). Melatonin treatment of pediatric residents for adaptation to night shift work. *Ambulatory Pediatrics*, 5(3), 172–177. <https://doi.org/10.1367/A04-124R.1>
- Costa, G. (2010). Shift work and health: Current problems and preventive actions. *Safety and Health at Work*, 1(2), 112–123. <https://doi.org/10.5491/SHAW.2010.1.2.112>
- Czeisler, C. A. (2010). Wake-promoting therapeutic medications not an appropriate alternative to implementation of safer work schedules for resident physicians. *Mayo Clinic Proceedings*, 85(3), 302–303. <https://doi.org/10.4065/MCP.2009.0723>
- Czeisler, C. A., Walsh, J. K., Wesnes, K. A., Arora, S., & Roth, T. (2009). Armodafinil for treatment of excessive sleepiness associated with shift work disorder: A randomized controlled study. *Mayo Clinic Proceedings*, 84(11), 958–972. <https://doi.org/10.4065/84.11.958>
- Farahmand, S., Vafaeian, M., Vahidi, E., Abdollahi, A., Bagheri-Hariri, S., & Dehpour, A. R. (2018). Comparison of exogenous melatonin versus placebo on sleep efficiency in emergency medicine residents working night shifts: A randomized trial. *World Journal of Emergency Medicine*, 9(4), 282–287. <https://doi.org/10.5847/wjem.j1920-8642.2018.04.008>
- Fletcher, K. E., Reed, D. A., & Arora, V. M. (2010). patient safety, resident education and resident well-being following implementation of the 2003 ACGME duty hour rules. *Journal of General*

- Internal Medicine*, 26(8), 907–919. <https://doi.org/10.1007/s11606-011-1657-1>
- FREIDA Online the AMA Residency and Fellowship Database AMA (2018). Retrieved from <https://www.ama-assn.org/life-career/search-ama-residency-fellowship-database>
- Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savovic, J., Schulz, K. F., Weeks, L., & Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, 343, d5928. <https://doi.org/10.1136/BMJ.D5928>
- Huffmyer, J. L., Kleiman, A. M., Moncrief, M., Scalzo, D. C., & Cox, D. J. (2019). Impact of caffeine ingestion on the driving performance of anesthesiology residents after 6 consecutive overnight work shifts. *Anesthesia and Analgesia*, 130(1), 66–75. <https://doi.org/10.1213/ANE.0000000000004252>
- Jackson, E. J., & Moreton, A. (2013). Safety during night shifts: A cross-sectional survey of junior doctors' preparation and practice. *British Medical Journal Open*, 3(9), e003567. <https://doi.org/10.1136/bmjopen-2013-003567>
- Jockovich, M., Cosentino, D., Cosentino, L., Wears, R. L., & Seaberg, D. C. (2000). Effect of exogenous melatonin on mood and sleep efficiency in emergency medicine residents working night shifts. *Academic Emergency Medicine*, 7(8), 955–958. <https://doi.org/10.1111/j.1553-2712.2000.tb02082.x>
- Kirsch, D. B., & Khosla, S. (2019). Public awareness, medical integration, and innovation in sleep medicine. *Journal of Clinical Sleep Medicine*, 15(05), 799–801. <https://doi.org/10.5664/jcsm.7786>
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of Internal Medicine*, 151(4), W-65–W-94. <https://doi.org/10.7326/0003-4819-151-4-200908180-00136>
- Liira, J., Verbeek, J. H., Costa, G., Driscoll, T. R., Sallinen, M., Isotalo, L. K., & Ruotsalainen, J. H. (2014). Pharmacological interventions for sleepiness and sleep disturbances caused by shift work. *Cochrane Database of Systematic Reviews*, (8), CD009776. <https://doi.org/10.1002/14651858.cd009776.pub2>
- McHill, A. W., Czeisler, C. A., & Shea, S. A. (2018). Resident physician extended work hours and burnout. *Sleep*, 41(8), zsy112. <https://doi.org/10.1093/sleep/zsy112>
- McKenna, H., & Wilkes, M. (2018). Optimising sleep for night shifts. *BMJ (Clinical Research Ed.)*, 360(March), j5637. <https://doi.org/10.1136/bmj.j5637>
- Mindell, J. A., Bartle, A., Wahab, N. A., Ahn, Y., Ramamurthy, M. B., Huong, H. T. D., Kohyama, J., Ruangdaraganon, N., Sekartini, R., Teng, A., & Goh, D. Y. T. (2011). Sleep education in medical school curriculum: A glimpse across countries. *Sleep Medicine*, 12(9), 928–931. <https://doi.org/10.1016/j.sleep.2011.07.001>
- Neil-Sztramko, S. E., Pahwa, M., Demers, P. A., & Gotay, C. C. (2014). Health-related interventions among night shift workers: A critical review of the literature. *Scandinavian Journal of Work, Environment and Health*, 40(6), 543–556. <https://doi.org/10.5271/sjweh.3445>
- Paez, A. (2017). Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*, 10(3), 233–240. <https://doi.org/10.1111/jebm.12266>
- Plesek, P. I. (2001). Redesigning health care with insights from the science of complex adaptive systems. In Institute of Medicine (US) Committee on Quality of Health Care in America (Ed.), *Crossing the quality chasm: A new health system for the 21st Century*. National Academies Press.
- Rose, S. H., & Curry, T. B. (2009). Mayo clinic proceedings. *Mayo Clinic Proceedings*, 84(11), 955–957.
- Rosenbaum, L., & Lamas, D. (2012). Residents' duty hours—toward an empirical narrative. *The New England Journal of Medicine*, 367(21), 2044–2049. <https://doi.org/10.1056/NEJMSr1210160>
- Ryan, R., & Hill, S. (2016). *How to GRADE the quality of the evidence*. Cochrane Consumers and Communication Group. Retrieved from <http://cccr.org/author-resources>. Version 3.0 December 2016.
- Schernhammer, E. S., Laden, F., Speizer, F. E., Willett, W. C., Hunter, D. J., Kawachi, I., Fuchs, C. S., & Colditz, G. A. (2003). Night-shift work and risk of colorectal cancer in the nurses' health study. *Journal of the National Cancer Institute*, 95(11), 825–828. <https://doi.org/10.1093/jnci/95.11.825>
- Shy, B. D., Phdc, I. P., & Nelson, L. S. (2011). Emergency medicine residents' use of psychostimulants and sedatives to aid in shift work. *American Journal of Emergency Medicine*, 29(9), 1034–1036.e1. <https://doi.org/10.1016/j.ajem.2010.06.004>
- Slanger, T. E., Gross, J. V., Pinger, A., Morfeld, P., Bellinger, M., Duhme, A.-L., Reichardt Ortega, R. A., Costa, G., Driscoll, T. R., Foster, R. G., Fritschi, L., Sallinen, M., Liira, J., & Erren, T. C. (2016). Person-directed, non-pharmacological interventions for sleepiness at work and sleep disturbances caused by shift work. *Cochrane Database of Systematic Reviews*, (8), CD010641. <https://doi.org/10.1002/14651858.CD010641.pub2>
- Smith-Coggins, R., Howard, S. K., Mac, D. T., Wang, C., Kwan, S., Rosekind, M. R., Sowb, Y., Balise, R., Levis, J., & Gaba, D. M. (2006). Improving alertness and performance in emergency department physicians and nurses: The use of planned naps. *Annals of Emergency Medicine*, 48(5), 596–604.e3. <https://doi.org/10.1016/j.annemergmed.2006.02.005>
- Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H. Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., ... Higgins, J. P. T. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*, 366, l4898. <https://doi.org/10.1136/bmj.l4898>
- Sugden, C., Housden, C. R., Aggarwal, R., Sahakian, B. J., & Darzi, A. (2012). Effect of pharmacological enhancement on the cognitive and clinical psychomotor performance of sleep-deprived doctors: A randomized controlled trial. *Annals of Surgery*, 255(2), 222–227. <https://doi.org/10.1097/SLA.0b013e3182306c99>
- Tempesta, D., Cipolli, C., Desideri, G., Gennaro, L. D., & Ferrara, M. (2013). Can taking a nap during a night shift counteract the impairment of executive skills in residents? *Medical Education*, 47(10), 1013–1021. <https://doi.org/10.1111/medu.12256>
- Trontell, M. C., Carson, J. L., Taragin, M. I., & Duff, A. (1991). The impact of the night float system on internal-medicine residency programs. *Journal of General Internal Medicine*, 6(5), 445–449. <https://doi.org/10.1007/BF02598169>
- Ulmer, C., Wolman, D. M., & Johns, M. M. E. (2009). Resident duty hours. Enhancing sleep, supervision, and safety. *Surgery*, 146(3), 398–409. <https://doi.org/10.1016/j.surg.2009.06.023>
- Wallach, S. L., Alam, K., Diaz, N., & Shine, D. (2006). How do internal medicine residency programs evaluate their resident float experiences? *Southern Medical Journal*, 99(9), 919–923. <https://doi.org/10.1097/01.smj.0000235501.62397.ef>
- Wang, F., Zhang, L., Zhang, Y., Zhang, B., He, Y., Xie, S., Li, M., Miao, X., Chan, E. Y. Y., Tang, J. L., Wong, M. C. S., Li, Z., Yu, I. T. S., & Tse, L. A. (2014). Meta-analysis on night shift work and risk of metabolic syndrome. *Obesity Reviews*, 15(9), 709–720. <https://doi.org/10.1111/obr.12194>
- Wegrzyn, L. R., Tamimi, R. M., Rosner, B. A., Brown, S. B., Stevens, R. G., Eliassen, A. H., Eliassen, A. H., Laden, F., Willett, W. C., Hankinson, S. E., & Schernhammer, E. S. (2017). Rotating night shift work and risk of breast cancer in the nurses' health studies. *American Journal of Epidemiology*, 186(5), 532–540. <https://doi.org/10.1093/aje/kwx140>

Weiss, P., Kryger, M., & Knauert, M. (2016). Impact of extended duty hours on medical trainees. *Sleep Health*, 2(4), 309–315. <https://doi.org/10.1016/j.sleh.2016.08.003>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Sholtes D, Kravitz HM, Deka A, Westrick J, Fogg LF, Gottlieb M. Optimising sleep and performance during night float: A systematic review of evidence and implications for graduate medical education trainees. *J Sleep Res.* 2020;00:e13212. <https://doi.org/10.1111/jsr.13212>